VŠB – Technická univerzita Ostrava
Fakulta elektrotechniky a informatiky
Katedra informatiky

Vážené varianty algoritmů pro určení centrality

Weighted Versions of Algorithms for Centrality Determination

2015 Martin Janek

VŠB - Technická univerzita Ostrava
Fakulta elektrotechniky a informatiky
Katedra informatiky

Zadání bakalářské práce

Student: **Martin Janek**

Studijní program: B2647 Informační a komunikační technologie

Studijní obor: 2612R025 Informatika a výpočetní technika

Téma: Vážené varianty algoritmů pro určení centrality
Weighted Versions of Algorithms for Centrality Determination

Zásady pro vypracování:

Zkoumání struktury a procesů v komplexních sítích reprezentujících reálné sítě různých typů (technické, informační, sociální apod.) je v současnosti dynamicky se vyvíjející oblastí. Určování různých typů centralit je jedním z častých testů pro analýzu komplexních sítí. Cílem práce je implementace algoritmů pro výpočet různých typů centralit pro komplexní sítě a jejich vážených variant.
1. Seznamte se s komplexními sítěmi a s vlastnostmi, které se nejčastěji zkoumají.
2. Seznamte se s centralitami a se způsobem jejich určení pro ohodnocené i neohodnocené (vážené i neválené) grafy.
3. Vyberte a naimplementujte algoritmy pro určení centralit.
4. Nad vybranými kolekcemi dat proveďte experimenty a jejich výsledky vhodně reprezentujte.

Seznam doporučené odborné literatury:

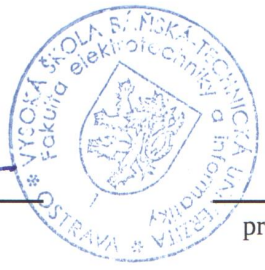
[1] M. E. J. Newman, Networks: An Introduction, Oxford University Press (2010), ISBN-10: 0199206651.
[2] http://toreopsahl.com/

D8le podle pokynů vedoucího bakalářské práce.

Formální náležitosti a rozsah bakalářské práce stanoví pokyny pro vypracování zveřejněné na webových stránkách fakulty.

Vedoucí bakalářské práce: **RNDr. Eliška Ochodková, Ph.D.**

Datum zadání: 01.09.2014
Datum odevzdání: 07.05.2015



doc. Dr. Ing. Eduard Sojka
vedoucí katedry

prof. RNDr. Václav Snášel, CSc.
děkan fakulty

Souhlasím se zveřejněním této bakalářské práce dle požadavků č. 26, odst. 9 *Studijního a zkušebního řádu pro studium v bakalářských programech VŠB-TU Ostrava.*

V Ostravě 7. května 2015 ..... Janek .....

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně. Uvedl jsem všechny prameny a publikace, ze kterých jsem čerpal.

V Ostravě 7. května 2015 ..... Janek .....

Rád bych na tomto místě poděkoval vedoucí práce RNDr. Elišce Ochodkové Ph.D., za její odborné rady, trpělivost, poskytnuté studijní materiály a čas strávený při konzultacích.

Abstrakt

Tato práce se zabývá tématikou centralit vrcholů v hranově ohodnocené komplexní síti. Přesněji se zabývá centralitami Degree, Closeness, Betwenness a PageRank.
V práci jsou popsány jednotlivé centrality a návrh a analýza implementace aplikace, schopné spočítat tyto centrality pro kolekce dat. Nakonec jsou v práci uvedené experimenty provedené nad kolekcemi dat.

Klíčová slova: komplexní síť, centralita, Degree, Closeness, Betweenness, PageRank

Abstract

This thesis deals with the theme of node centrality in weighted network. Strictly deals with Degree, Closeness, Betweenness and PageRank centrality.
This thesis describes the individual centrality and design and analysis of application implementation, which is able to calculate these centrality on collections of data. Finally, this thesis mentions experiments carried out on collections of data.

Keywords: Complex network, Centrality, Degree, Closeness, Betweenness, PageRank

Seznam použitých zkratek a symbolů

CSV	–	Comma Separated Values
URL	–	Uniform Resource Locator

Obsah

Seznam tabulek

Seznam obrázků

1 Úvod

V posledních letech se díky dostupnosti dat na internetu a nárůstu vypočetního výkonu počítačů stala problematika analýzy komplexních sítí velmi oblíbenou. Příklady systémů, které můžeme pomocí komplexních sítí reprezentovat, lze v reálném světě najít velmi mnoho. Jedná se například o sociální sítě, citační sítě, dopravní sítě, elektrické rozvodné sítě, nervové sítě v biologických systémech a mnohé jiné.

Komplexní sítě se matematicky reprezentují pomocí grafů, kde jednotlivé objekty sítě představují vrcholy grafu a definovaný typ interakce mezi objekty sítě je reprezentován hranami grafu. Ve srovnání s jednoduchými grafy, jako jsou např. mřížky nebo náhodné grafy, se komplexní sítě liší svými netriviálními vlastnostmi.

Komplexní sítě jsou velmi rozsáhlé struktury, často až s milióny vrcholů a hran. Právě tato rozsáhlost komplikuje jejich analýzu. Vědci se zabývají zkoumáním různých vlastností sítí, včetně topologie, která je velmi důležitá, protože struktura ovlivňuje procesy v síti. Např. topologie sociálních sítí ovlivňuje rychlost šíření informací, nebo nemocí a můžeme z nich vyčíst, kteří jednotlivci mají v síti největší vliv. Pokud tedy pochopíme struktutu sítě, můžeme odhadnout její chování.

Mezi vlastnosti komplexních sítí patří také centrality určující důležitost jednotlivých objektů v rámci celé sítě. V této práci se budu věnovat čtyřem druhům centralit. Jsou to degree, closeness, betweenness a pageRank.

Cílem této práce je implementace aplikace schopné načíst kolekci dat, spočítat vybrané centrality pro hranově ohodnocené i neohodnocené grafy. Aplikace bude schopna zapsat vypočtená data do souboru, seřazená vzestupně.

V první kapitole této práce jsou vysvětleny komplexní sítě a jednotlivé centrality. V druhé kapitole je popsáný návrh a analýza implementace aplikace. V třetí kapitole jsou uvedeny experimenty se získanými daty.

2 Komplexní sítě

V této kapitole jsou popsány komplexní sítě a vybrané typy centralit implementovaných v aplikaci [9].

2.1 Co je síť?

Síť je kolekce entit, které jsou propojeny vazbami. Například lidé, a jejich přátelé, počítače propojené společnou počítačovou síti nebo webové stránky, odkazující na jiné webové stránky.

2.2 Dělení komplexních sítí

V této kapitole jsou uvedené některé typy komplexních sítí, které se od sebe rozlišují podle toho, co představují jejich uzly, jejich hrany a v případě ohodnocených hran také podle toho, co představuje váha hrany.

2.2.1 Sociální sítě

Prvním typem komplexní sítě je sociální síť. Příkladem takové sociální sítě je Facebook, kde jednotlivé uzly představují samotné uživatelé a hrany mezi uzly mohou představovat přátelství uživatelů, jinak řečeno, dva uživatelé jsou přáteli. Síť uživatelů a jejich vzájemné vazby jsou zobrazeny na obrázku 1 [17]. V případě hranově ohodnocené sociální sítě může hrana mezi dvěma uzly, tedy přátelství mezi dvěma uživateli, představovat počet společných přátel nebo počet vzájemně odeslaných zpráv. Hrana také může představovat fotografii, na které jsou oba uživatelé označení.

Příkladem sociální sítě je také e-mailová síť. V tomto případě jednotlivé uzly představují uživatelské účty, resp. e-mailové adresy a hrany mezi uzly představují odeslané e-mailové zprávy mezi uživatelskými účty. Pokud je síť hranově ohodnocená, mohou váhy hran mezi uzly představovat počet odeslaných e-mailových zpráv mezi uživatelskými účty.

Dalšími příklady sociálních sítí jsou herecké sítě a sítě autorů, resp. spoluautorů.

2.2.2 Informační sítě

Informační síť je dalším typem komplexní sítě. Účelem takové sítě je sdružovat (spojovat) informace.

Informační sítí je například citační síť. Různé diplomové práce, vědecké publikace, články a texty mají jako své zdroje uvedené jiné vědecké práce, publikace či diplomové práce. Přesněji řečeno, uzly sítě představují samotné texty. Pokud mezi dvěma uzly existuje hrana, znamená to, že se jeden text odkazuje na druhý text ve svých zdrojích či referencích. Váha hrany může znamenat počet odkazů mezi texty.

Mezi informační sítě patří také The World Wide Web. Uzly této sítě jsou reprezentovány jako webové stránky a hrana z jednoho uzlu do druhého je reprezentována jako odkaz na jinou webovou stránku.

Obrázek 1: Sociální síť

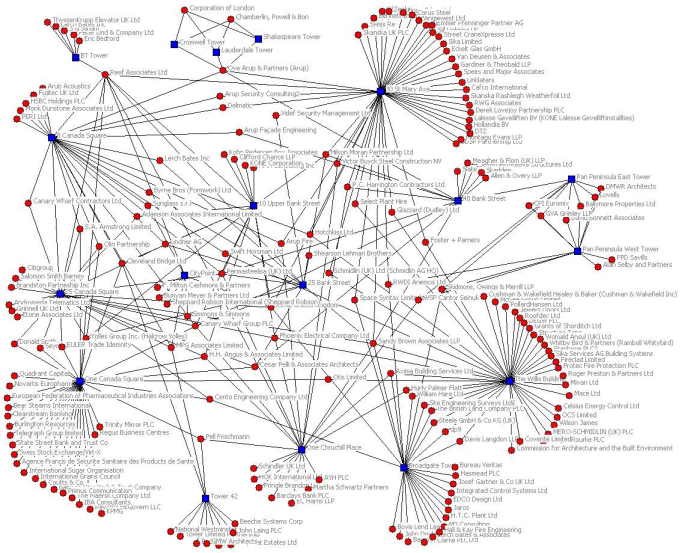
2.2.3 Technologické sítě

Technologické sítě jsou velmi rozsáhlé, jsou vybudované pro účely distribuce zboží. Takovou celosvětovou technologickou sítí je Internet. Jeho uzly jsou servery, které jsou rozmístěné po celém světě. Hrany sítě představují propojení těchto serverů, např. kabelem. Váha hrany může představovat délku kabelu, nebo rychlost přenosu informací. Celosvětové internetové propojení lze vidět na obrázku 2 [18].

Obrázek 2: Internet

Obchodní síť je další celosvětově rozsáhlou sítí, jejíž uzly jsou jednotlivé firmy a hrany představují obchodní kontrakty mezi firmami. Tato síť může být i hranově ohodnocená, kdy váha hrany může představovat počet obchodních kontraktů mezi firmami, nebo množství nakoupeného zboží. Příklad obchodní sítě lze vidět na obrázku 3 [19].

Technologickými sítěmi jsou také transportní sítě. Mezi ně se řadí např. silniční, železniční a energetické sítě. V těchno sítích uzly představují dopravní uzly - křižovatky, nádraží, elektrické rozvodny a hrany jsou cesty mezi těmito uzly, např. silnice, železnice,



Obrázek 3: Obchodní síť

nebo dráty vysokého napětí. Váha takových hran může znamenat vzdálenost mezi uzly, tedy délku silnice mezi dvěma křižovatkama, maximální povolenou rychlost, hustotu provozu na komunikacích, počet přepravených pasažérů a další.

Další sítí tohoto druhu je telefonní síť. Uzly této sítě mohou představovat telefonní ústředny, které jsou vzájemně propojené kabely, tedy hranami. Váha těchto hran by mohla nést informaci o délce kabelu. Uzly telefonní sítě by mohly také představovat samotné telefony, tedy telefonní čísla a hrany mezi těmito uzly by znamenaly hovory a váhy těchto hran by zaznamenávaly počet hovorů nebo počet provolaných minut.

Mezi technologické sítě patří také sítě aerolinií. Tyto sítě zaznamenávají letové dráhy letadel mezi jednotlivými letišti. Je patrné, že v tomto případě jsou uzly sítě reprezentované jako letiště a hrany sítě jsou reprezentované jako letové dráhy (letové linky) mezi letišti. Takové sítě mohou být velmi rozsáhlé, neboť mohou zaznamenávat letové trasy po celém světě, jak lze vidět na obrázku 4 [20]. Váha ohodnocených hran sítě aerolinií může mít mnoho významů. Například může představovat letovou vzdálenost, počet sedadel, které jsou v letadle k dispozici, počet převážených pasažérů, nebo počet provedených letů.

2.2.4 Biologické sítě

Biologické sítě jsou také hojně využívané. Zobrazují například potravinové řetězce, metabolismus, interakce proteinů, nervová síť, synaptická síť mozku a mnoho dalších. Obrázek 5 [21] zobrazuje mozkovou síť neuronů. V potravinovém řetězci mohou být živočišné druhy znázorněné jako uzly a pokud je mezi dvěma uzly hrana, znamená to, že jedno zvíře žere druhé.

Obrázek 4: Síť letových tras

2.3 Graf

Jelikož se komplexní sítě modelují jako grafy, tak v této práci místo pojmu komplexní síť uvádím pojem graf, místo pojmu uzel komplexní sítě uvádím pojem vrchol grafu.

Pojem grafu byl zaveden Leonhardem Eulerem v roce 1736. Jedná se o model, který reprezentuje objekty a vztahy mezi nimi. Graf se dá reprezentovat jako obrázek objektů či míst a jejich vzájemného propojení na základě nějakého vztahu. Například objekty mohou být lidé a propojení může být na základě příbuzenského vztahu [10].

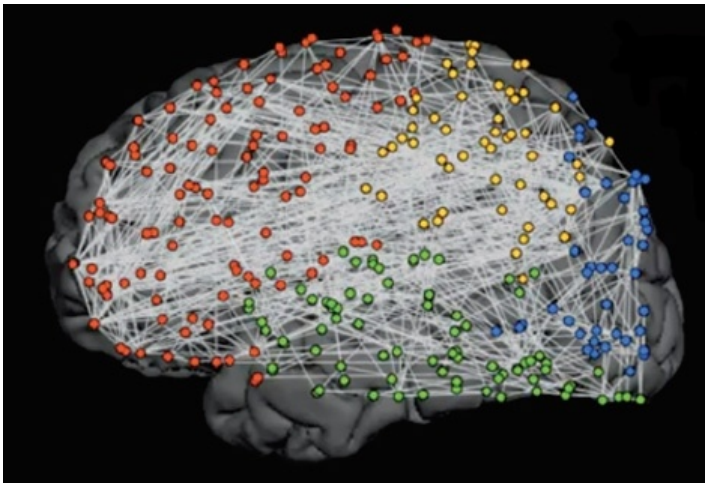
Definice 1 *Graf G je uspořádaná dvojice G = (V, E), kde V je neprázdná množina vrcholů a E je množina hran - množina dvouprvkových podmnožin množiny V.*

Takto definovaný graf se nazývá *neorientovaný graf*, u kterého se nerozlišují hrany uv a vu. Příkladem takového grafu může být potrubní síť s čerpadlem, kde kapalina může téct oběma směry nebo dopravní síť dálnic, po kterých se dá jezdit oběma směry.

Definice 2 *Neorientovaný graf G je dvojice (V, E), kde V je konečná množina objektů, kterým říkáme vrcholy, někdy též uzly grafu a E je množina některých dvojic vrcholů, kterým říkáme hrany grafu.*

Při řešení některých problémů má smysl rozlišit pořadí vrcholů hrany. V takovém případě se jedná o *orientovaný graf*.

Orientované grafy se vyskytují při řešení různých praktických problémů. Příkladem může být elektrická síť modelovaná grafem, ve které se dopravuje energie od zdroje ke



Obrázek 5: Síť neuronů v mozku

spotřebičům, silniční síť, kde se mohou vyskytovat jednosměrky nebo potrubní síť bez čerpadel, kde kapalina teče vždy jedním směrem.

Definice 3 *Orientovaný graf G je dvojice (V, E), kde V je konečná množina objektů, kterým říkáme vrcholy, někdy též uzly grafu a E je množina některých uspořádaných dvojic vrcholů, kterým říkáme orientované hrany grafu. Orientace hrany bývá v grafu zobrazena šipkou.*

Stupeň vrcholu v jednoduchého grafu je roven součtu hran incidentních s vrcholem. U orientovaného grafu je celkový stupeň vrcholu v vždy roven součtu odchozího a příchozího stupně vrcholu v.

Definice 4 *Mějme neprázdnou množinu vrcholů V. Dvojice (V,A) je nějaká podmnožina kartézského součinu V x V, se nazývá orientovaný graf D nebo stručně digraf D. Prvkům množiny A říkáme orientované hrany. Jsou-li u, v dva různé vrcholy digrafu D, tak orientovanou hranu (u,v) budeme značit stručně uv. Vrchol u se nazývá výchozí nebo počáteční a v koncový vrchol hrany uv. Hrany uv a vu z množiny A(D) se nazývají opačně orientované nebo jen opačné hrany.*

Grafy se rozdělují na dva základní druhy podle toho, zda jsou jeho hrany orientované.

2.4 Vlastnosti sítí

Definice 5 *Cesta v grafu je posloupnost vrcholů, pro kterou platí, že v grafu existuje hrana z daného vrcholu do jeho následníka. Žádné dva vrcholy a žádné dvě hrany se přitom neopakují.*

Mezi libovolnými dvěma vrcholy u a v může existovat více cest různých délek. Cesta z vrcholu u do vrcholu v, která má nejmenší vzdálenost se nazývá nejkratší cesta. Pokud však mezi vrcholy u a v neexistuje žádná hrana, pak nejkratší cesta mezi těmito vrcholy je rovna ∞.

Definice 6 *Průměr grafu je nejdelší vzdálenost kroků mezi libovolnými dvěma uzly.*

- Průměr grafu je někdy nazývaný také jako excentricita grafu. Jestliže průměr grafu *G* je roven číslu *k*, tak musí v grafu *G* existovat dvojice vrcholů *u* a *v*, že jejich vzdálenost, tedy délka nejkratší cesty *(u,v)* je *k*. Takové cestě se říká *diametrická cesta*. Čím více hran je v grafu, tím menší je průměr grafu a naopak grafy s velkým průměrem jsou řídké.

  Definice 7 *Největší excentricita v grafu G se nazývá průměr grafu a nejmenší excetricita se nazývá poloměr grafu G. Průměr grafu G se značí diam(G) a poloměr rad(G).*

- Stupeň je vlastnost vrcholu grafu, která udává počet hran, které do daného vrcholu zasahují.

  Definice 8 *Stupeň vrcholu v je počet hran, se kterými je vrchol v incidentní, značí se deg(v).*

  V orientovaném grafu jsou hrany orientované, proto se rozlišuje vstupní stupeň $deg^+(v)$, počet hran, které vcházejí do vrcholu v a výstupní stupeň $deg^-(v)$, počet hran, které vycházejí z vrcholu v. Celkový stupeň vrcholu je pak roven součtu vstupního a výstupního stupně.

- Shluky v grafu vznikají tehdy, je-li uzel a spojen hranou s uzlem b a zároveň uzel b je spojen s uzlem c, pak je pravděpodobné, že uzel a bude spojen s uzlem c.

2.5 Centrality

Centralita vrcholu je jedna z metrik zachycující jednotlivé vlastnosti topologie sítě. Centralita určuje, jak je vrchol v rámci sítě důležitý. Určuje, které vrcholy jsou nejdůležitější nebo centrální v síti z hlediska struktury sítě [2].

První zmínku o pojmu *centralita*, pocházející z oblasti sociologie, můžeme najít v [3], kde autor definuje sadu metod centrality, založených na betweenness. Freeman ve své práci definoval centrality Degree, Closeness, Betweenness a Eigenvector [11], [12].

V této práci se zabývám centralitami Degree, Closeness, Betweenness a PageRank.

2.5.1 Degree

Degree centralita je nejjednodušší mírou centrality. Určuje počet hran nebo součet vah hran incidentních s vrcholem, tedy počet přímých vazeb k dalším vrcholům. V hranově ohodnoceném grafu s orientovanými hranami se rozlišují Degree hodnoty vrcholu vstupující (in-Degree) a Degree hodnoty vrcholu vystupující (out-Degree), podle toho, jestli hrany do vrcholu vstupují nebo z něj vystupují. Hodnota in-Degree vrcholu v je tedy součet vah hran vstupujících do vrcholu v a hodnota out-Degree je součet vah hran vystupujících z vrcholu v. Celková hodnota Degree centrality vrcholu v je pak rovna podílu in-Degree/out-Degree.

Vrchol s vysokým počtem hran nebo více spojeními je ve struktuře grafu více centrální a má tak větší schopnost ovlivňovat ostatní vrcholy. Vrchol, který má vysoké in-Degree, je v rámci grafu velmi populární a vrchol, který má vysoké out-Degree je v rámci grafu velmi vlivný a má vysokou šanci ovlivnit ostatní.

V sociální síti, čím větší je stupeň vrcholu, tím vlivnější je člověk. V citační síti, čím větší má vrchol in-Degree, tím větší vliv měla publikace na vědecký výzkum.

Aby bylo možné porovnat vrcholy různých grafů podle hodnot Degree centrality, je nutné tyto hodnoty nejprve normalizovat. Normalizace Degree hodnot spočívá v tom, že se hodnota Degree centrality každého vrcholu vydělí maximálním počtem všech hran, které vrchol může mít, tj. *(n-1)*, kde *n* je počet všech vrcholů grafu [4], [5].

Obecný vzorec pro výpočet Degree centrality jednotlivých vrcholů neorientovaného a neohodnoceného grafu, viz vzorec 1, kde C_D je hodnota Degree centrality vrcholu u, $d(u)$ je stupeň vrcholu u a N je celkový počet vrcholů v grafu.

$$C_D(u) = \frac{d(u)}{N-1} \tag{1}$$

Vzorec 2 znázorňuje výpočet jednotlivých vrcholů orientovaného a ohodnoceného grafu, kde $d_{in}(u)$ je součet vah vstupních hran vrcholu u a $d_{out}(u)$ je součet vah výstupních hran vrcholu u.

$$C_D(u) = \frac{d_{in}(u)/d_{out}(u)}{N-1} \tag{2}$$

2.5.2 Closeness

Closeness centralita, neboli blízkost vrcholu, udává informaci o tom, jak dlouho bude trvat, než se určitá informace rozšíří z daného vrcholu do všech ostatních vrcholů grafu.

Matematicky lze hodnotu Closeness centrality vypočítat jako převrácenou hodnotu součtu nejkratších cest ke všem ostatním vrcholům. Pro nalezení minimálních vzdáleností lze použít např. Dijkstrův algoritmus nebo Floyd-Warshallův algoritmus.

Closeness centralitu lze počítat pouze pro souvislé grafy, tzn. že se lze z libovolného vrcholu grafu dostat po hranách ke všem ostatním vrcholům grafu. Pokud tomu tak není, vypočítají se hodnoty Closeness centrality vrcholů v každé části grafu zvlášť a poté se tyto hodnoty normalizují. Normalizace Closeness hodnoty vrcholu v se provede tak, že se jeho Closeness hodnota vynásobí počtem všech vrcholů té části grafu, ve které se vrchol v nachází [4]. Po provedené normalizaci hodnot vrcholů lze porovnávat různé grafy podle Closeness centrality.

Vzorcem Closeness centrality je vzorec (3), kde $C_c(u)$ je hodnota closeness centrality vrcholu u, V je množina všech vrcholů grafu a $d(u,v)$ je délka nejkratší cesty z vrcholu u do vrcholu v.

$$C_c(u) = \sum_{v \in V} \frac{1}{d(u,v)} \tag{3}$$

Jestliže je graf ohodnocený a váhy hran představují např. vzdálenost, tzn. čím vyšší je váha hrany (u, v), tím větší je vzdálenost mezi vrcholy u a v, tak se výpočet nemění. Pokud však váhy hran představují např. blízkost, tzn. čím větší je váha hrany (u, v), tím blíže jsou vrcholy u a v k sobě, tak se ve výpočtu vzdálenosti $d(u, v)$ musí počítat s obrácenými hodnotami vah hran, tj. $1/d(u, v)$, nebo jednodušší způsob je $(C_c(u))^{-1}$.

2.5.3 Betweeness

Betweenness centralita určuje, kolik cest mezi dvojicemi ostatních vrcholů prochází právě daným vrcholem. Vrchol s vysokou hodnotou Betweenness centrality má významnou roli v propojování odlišných skupin. Příkladem je graf, jehož vrcholy reprezentují osoby, které navštěvují dva různé zájmové kroužky. Každá osoba navštěvuje pouze jeden zájmový kroužek, kromě jedné jediné osoby (označená jako osoba X), která navštěvuje oba dva kroužky. Tato osoba X jako jediná propojuje obě skupiny osob a tudíž má největší hodnotu Betweenness centrality ze všech. Tato osoba X bude značně ovlivňovat dění v grafu např. blokováním nežádoucích zpráv, vybíráním poplatků za spojení nebo izolováním některé osoby, která nemá jinou možnost jak se ke sdílené informaci dostat.

Pro výpočet Betweenness centrality vrcholu u je důležité znát, kolik nejkratších cest mezi dvěma libovolnými vrcholy prochází přes vrchol u. Výpočet znázorňuje vzorec (4), kde $C_B(u)$ je hodnota Betweenness centrality vrcholu u, $g_{j,u,k}$ je počet nejkratších cest mezi vrcholy j a k, které procházejí přes vrchol u a $g_{j,k}$ je počet všech nejkratších cest mezi vrcholy j a k.

$$C_B(u) = \sum_{j,k \in V; j,k \neq u} \frac{g_{j,u,k}}{g_{j,k}} \tag{4}$$

Aby bylo možné porovnávat různé grafy podle Betweenness centrality, je nutné normalizovat hodnoty Betweenness centrality všech vrcholů grafu. Normalizace Betweenness hodnoty vrcholu v spočívá ve vydělení hodnoty Betweenness centrality vrcholu v počtem všech možných hran, které by graf mohl obsahovat, tj. $(n-1) * (n-2)$ pro orientovaný graf a $(n-1) * (n-2)/2$ pro neorientovaný graf, kde n je počet vrcholů grafu [4].

Pokud je graf ohodnocený, je nutné při zjišťování nejkratších cest mezi dvěma vrcholy brát v úvahu váhy hran.

Výše uvedená varianta výpočtu Betwenness centrality využívá pouze nejkratší cesty mezi dvěma vrcholy. Existuje ale i další způsob výpočtu Betweenness centrality, který využívá všechny cesty mezi dvěma vrcholy. To znamená, že při komunikaci mezi dvěma vrcholy nemusí být použity pouze nejkratší cesty. Je-li nejkratší cesta nedostupná, použije se druhá nejkratší cesta, samozřejmě pokud nějaká druhá existuje [6]. Tento způsob výpočtu se ale moc nepoužívá, protože je výpočetně náročný.

2.5.4 PageRank

PageRank centralita udává významnost vrcholu grafu v závislosti na propojení s jinými vrcholy. Algoritmus pro výpočet PageRank centrality využívají internetové vyhledávače,

např. Google.com. Algoritmus určuje významnost webových stránek, které se využívají při řazení nalezených stránek v internetovém vyhledávači. Např. při určování důležitosti webové stránky v se využívají hypertextové odkazy, které odkazují na webovou stránku v z jiných webových stránek a významnosti těchto stránek. V grafu komplexní sítě představují webové stránky vrcholy grafu, hrany grafu představují hypertextový odkaz z jedné stránky na jinou a váhy hran znamenají počet odkazů na stránku.

Hodnota PageRank centrality webové stránky udává pravděpodobnost, s jakou se dá dostat díky hypertextovým odkazům na danou webovou stránku. Součet hodnot PageRank centrality všech vrcholů grafu je roven 1, tedy 100% pravděpodobnost.

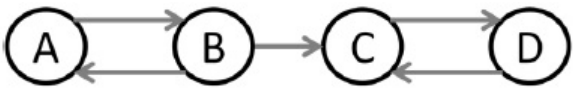
Vzorec (5) popisuje algoritmus výpočtu PageRank centrality, kde PR_{x+1} je hodnota PageRank centrality vrcholu A v iteraci s x, U je množina všech vrcholů odkazujících na vrchol A a N_u je počet výstupních hran vrcholu u.

$$PR_{x+1}(A) = \sum_{u \in U} \frac{PR_x(u)}{N_u} \tag{5}$$

Vzorec (5) ale neřeší problém slepých vrcholů, tj. vrcholů, bez jakýchkoliv výstupních hran. U těchto vrcholů se ztrácí hodnoty PageRank centrality a součet hodnot PageRank centrality všech vrcholů grafu pak přestává být roven 1. Nejčastějším řešením je, každému slepému vrcholu přidat výstupní hrany vedoucí na všechny vrcholy grafu i na sebe sama. Tyto nově přidané výstupní hrany se nemusí přímo doplňovat do grafu, stačí s nimi jen počítat, jak je ukázáno ve vzorci (6), kde D je množina všech slepých vrcholů grafu, V je množina všech vrcholů grafu a $|V|$ je velikost množiny $|V|$.

$$PR_{x+1}(A) = \sum_{u \in U} \frac{PR_x(u)}{N_u} + \frac{1}{|V|} \sum_{s \in D} PR_x(s) \tag{6}$$

Dalším problémem, který je třeba vyřešit se nazývá *Rank sink* (zaniknutí hodnoty), který vzniká, pokud vrcholy v jedné skupině odkazují samy na sebe, ale neodkazují na vrcholy v jiné skupině a zároveň je skupina odkazovaná z vnějšku z jiné skupiny. Problém je vidět na obrázku 6, kde vrcholy A a B při výpočtu PageRank centrality postupně předají své PageRank hodnoty vrcholům C a D. PageRank hodnoty vrcholů A a B budou pak rovny nule. Navíc může dojít k dalšímu problému, když nebudou mít vrcholy C a D každý přesně polovinu celkového hodnoty PageRank centrality, tak si vrcholy budou v každé iteraci vzájemně vyměňovat své PageRank hodnoty a nikdy nenastane ustálený stav, výpočet se bude pořád do nekonečna opakovat.



Obrázek 6: Klesání hodnoty při výpočtu PageRanku.

Problém *Rank sink* se dá vyřešit doplněním vzorce 6 o konstantu nazvanou faktor tlumení. Vysvětlit se to dá na příkladu, kdy se nějaká osoba pohybuje Webem prostřednictvím hypertextových odkazů nebo využívá tzv. teleport tak, že přejde na náhodnou

stránku zadáním její URL do webového prohlížeče. Takže osoba s pravděpodobností d následuje hypertextové odkazy nebo s pravděpodobností $(d-1)$ využívá teleport. Faktor tlumení má obvykle nastavenou hodnotu na 0.85, ale tato hodnota může být měněna. Čím blíže je hodnota faktoru tlumení jedné, tím více iterací je potřeba k dosažení ustáleného stavu [7].

Vzorec (7) zobrazuje algoritmus pro výpočet PageRank centrality doplněný o faktor tlumení, následnou normalizaci získané PageRank hodnoty, aby bylo možné porovnávat různé grafy podle PageRank centrality a váhy hran, aby bylo možné počítat PageRank centralitu pro ohodnocené grafy. Proměnná w_{utoA} je váha hrany vedoucí z vrcholu u do vrcholu A a w_{out} je součet vah všech výstupních hran vrcholu u.

$$PR_{x+1}(A) = \frac{(1-d)}{|V|} + d \cdot (\sum_{u \in U} \frac{PR_x(u) \cdot w_{utoA}}{w_{out}} + \frac{1}{|V|} \sum_{s \in D} PR_x(s)) \tag{7}$$

3 Analýza a návrh aplikace

Tato kapitola je zaměřena na tvorbu aplikace z pohledu softwarového inženýra. Je zde rozebrán návrh tříd a je zde stručně popsáno, k čemu každá třída slouží a jakou v aplikaci hraje roli. Také je zde popsáno uživatelské rozhranní aplikace a ovládání aplikace.

3.1 Zadání aplikace

Aplikace bude sloužit k výpočtu centralit uzlů pro hranově ohodnocené grafy.

Aplikace bude schopna načíst kolekci dat ze souboru typu CSV. Kolekcí dat se myslí komplexní síť (graf) tvořená vrcholy a ohodnocenými hranami spojujicími vrcholy. Struktura vstupního CSV souboru bude taková, že na každém řádku bude jedna hrana, reprezentována Id prvního vrcholu, Id druhého vrcholu a váhy hrany. Všechny tři hodnoty budou od sebe oddělené středníkem.

Pro každý vrchol v kolekci se spočítají čtyři typy centralit pro ohodnocené i neohodnocené grafy. Bude se také měřit doba výpočtu centralit. Aplikace bude schopna přepsat váhy všech hran na 1 a tím změnit ohodnocený graf na neohodnocený.

Vypočtená data se budou zapisovat do souboru typu CSV. Pro každou datovou kolekci a každou centralitu bude vytvořen samostatný CSV soubor. Struktura výstupního CSV souboru bude Id vrcholu a jeho hodnota centrality, obě hodnoty budou oddělné středníkem. Jeden vrchol na jednom řádku.

3.2 Návrh tříd

Na obrázku 7 lze vidět diagram tříd.

Třída *Vertex* představuje vrchol grafu. Každá instance této třídy má své Id, kolekci svých vstupních hran a hodnoty čtyř centralit.

Třída *Edge* představuje hranu grafu. Každá hrana obsahuje své Id, dále Id dvou vrcholů, které daná hrana spojuje a svou váhu. Díky této váze bude aplikace schopna spočítat varianty centralit pro hranově ohodnocené grafy. Pokud bychom chtěli z hranově ohodnoceného grafu udělat hranově neohodnocený, stačí všem hranám grafu přepsat jejich váhu na 1.

Nejdůležitější třídou je třída *Graph*, která představuje celou kolekci dat, tedy celou komplexní síť. Tato třída obsahuje kolekci vrcholů a kolekci hran nacházejicích se v grafu. Třída bude obsahovat funkci pro vyhledání vrcholu grafu podle jeho Id, funkci pro vyhledání hrany grafu podle Id dvou vrcholů incidentních s danou hranou, funkci, která bude vracet váhu hrany, kterou nalezne podle Id dvou vrcholů dané hrany zadaných jako parametry funkce. Dále bude obsahovat funkce, které spočítají sumy váh výstupních a vstupních hran vrcholu, jehož Id bude zadané parametrem fukce. Další funkce bude sloužit k přeměnění ohodnoceného grafu na neohodnocený tím, že přepíše váhy všech hran na 1.

Třída *LoadData* bude zprostředkovávat otevření vstupního souboru. Z tohoto souboru pak načte a zpracuje data. Z těchto dat poté vytvoří instanci třídy *Graph*, která bude reprezentovat celou datovou kolekci.

Třída *ExportData* bude mít za úkol z dané instance třídy *Graph* získat data a zapsat je do souboru typu CSV.

Třídy *Degree*, *Closeness*, *Betweenness* a *PageRank* budou mít za úkol získat z instance třídy *Graph* potřebná data k výpočtu ohodnocených i neohodnocených variant centralit Degree, Closeness, Betweenness a PageRank. Vypočtené hodnoty centralit každého vrcholu grafu budou uloženy do proměnných instance daného vrcholu (instance třídy *vertex*).

Třída *BubbleSort* se bude sloužit k setřídění výstupních dat vzestupně podle hodnoty centrality. Setřídění bude použito v třídě *ExportData* pro zápis vypočtených hodnot centralit do souboru. Třída FloydWarshall bude sloužit k nalezení nejkratších cest mezi všemi dvojicemi vrcholů v grafu. Instance této třídy bude obsažena ve třídách *Closeness* a *Betweenness*, ve kterých je nalezení nejkratších cest důležité pro výpočet Closeness a Betweenness centralit.

Podrobný pohled na třídy, jejich atributy a metody je na obrázku 8

3.3 Uživatelské rozhraní

Tato aplikace, původně plánovaná jako konzolová aplikace, byla vytvořena ve vývojovém prostředí Microsoft visual Studio 2013, ve kterém bylo k aplikaci vytvořeno také uživatelské rozhraní (dále jen jako UR), které lze vidět na obrázku 9.

Součástí UR je DropDownList, do kterého jsou při inicializaci aplikace vloženy názvy všech souborů, které se nacházejí v adresáři pro vstupní kolekce dat. Dále jsou v UR obsažené tři tlačítka pro načtení kolekce dat z CSV souboru, pro výpočet centralit a pro zápis vypočtených hodnot do CSV souboru. Název CSV souboru bude složen následovně: [název datové kolekce]_[weighted/unweighted]_[název centrality] .csv. Dále je v UR pět checkboxů. Čtyři z nich slouží k aktivaci výpočtu čtyř typů centralit a pátý checkbox nastavuje, zda se budou počítat ohodnocené nebo neohodnocené varianty centralit. Vedle checkboxu pro PageRank centralitu se nachází numerický vstup, který nastavuje hodnotu pro dampending factor, neboli faktor útlumu, který byl zmíněný v kapitole 2.5.4 na straně 12. Ve spodní části UR je informační label, do kterého se vypisují oznámení o běhu aplikace, např. že data byla úspěšně načtená ze souboru nebo že data byla úspěšně zapsaná do souboru. Pokud dojde při běhu aplikace k nějaké chybě, vypíše se chybové oznámení do labelu.

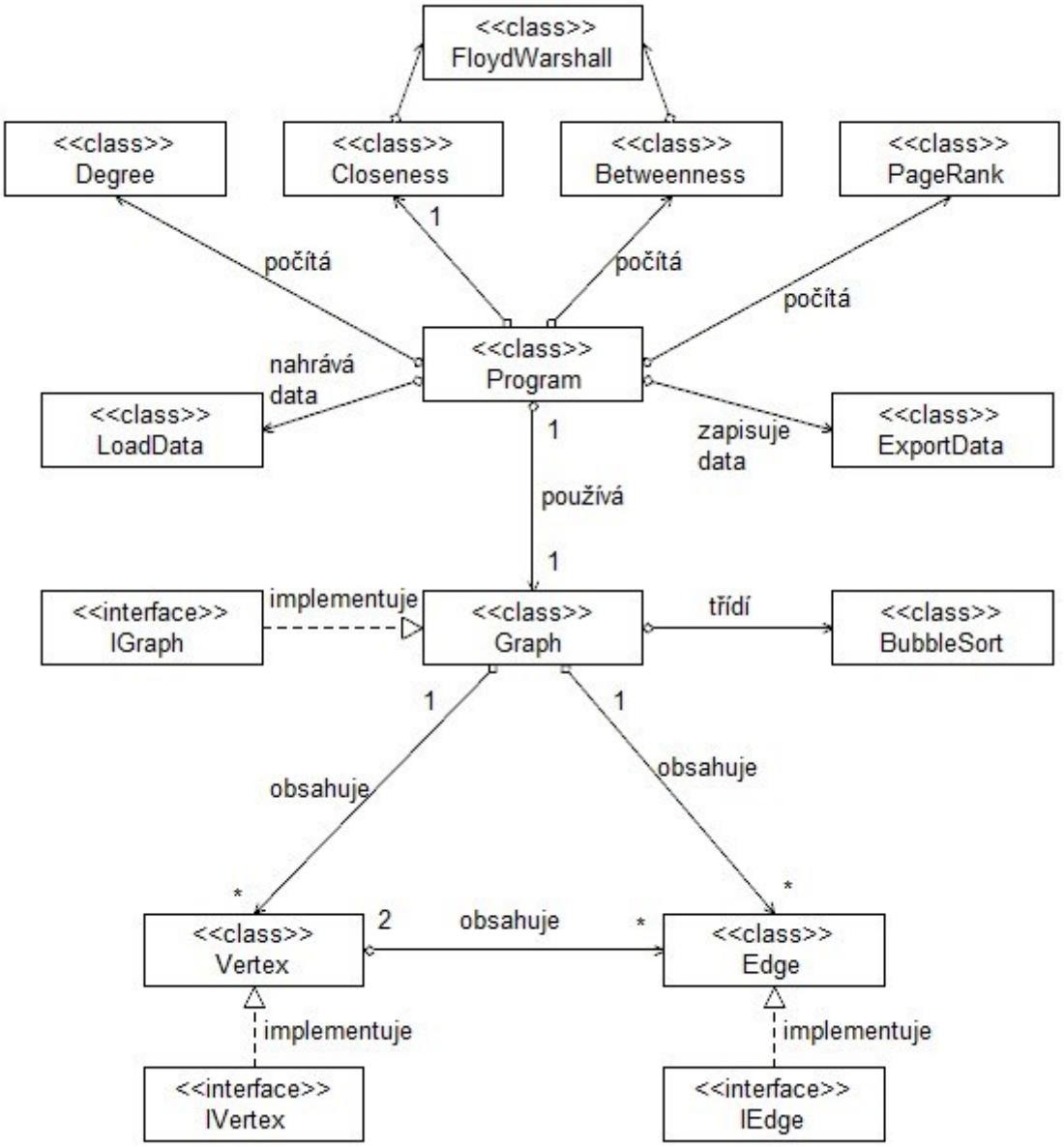
Postup k ovládání aplikace je následovný:

1. Vybrat v dropDownListu název datové kolekce, kterou chceme zpracovat a zmáčknout tlačítko "Load". Pokud vše proběhne v pořádku, vypíše se do info.labelu informace o datové kolekci, jako jsou např. název datové kolekce, počet vrcholů a hran, apod.

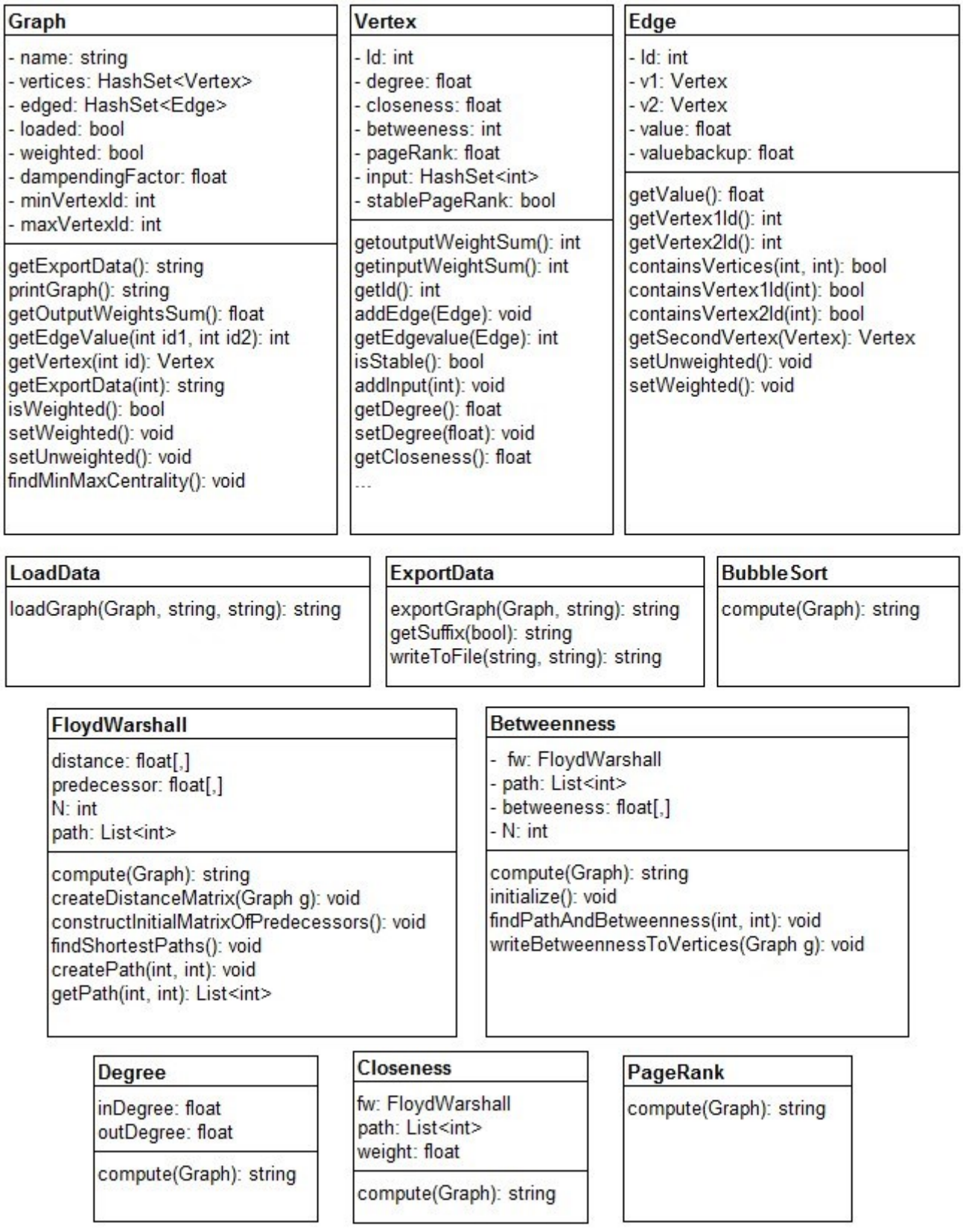
2. Označit centrality, které chceme počítat.

3. Pokud je potřeba, upravit hodnotu faktoru útlumu a jeho změnu potvrdit stisknutím klávesy Enter.

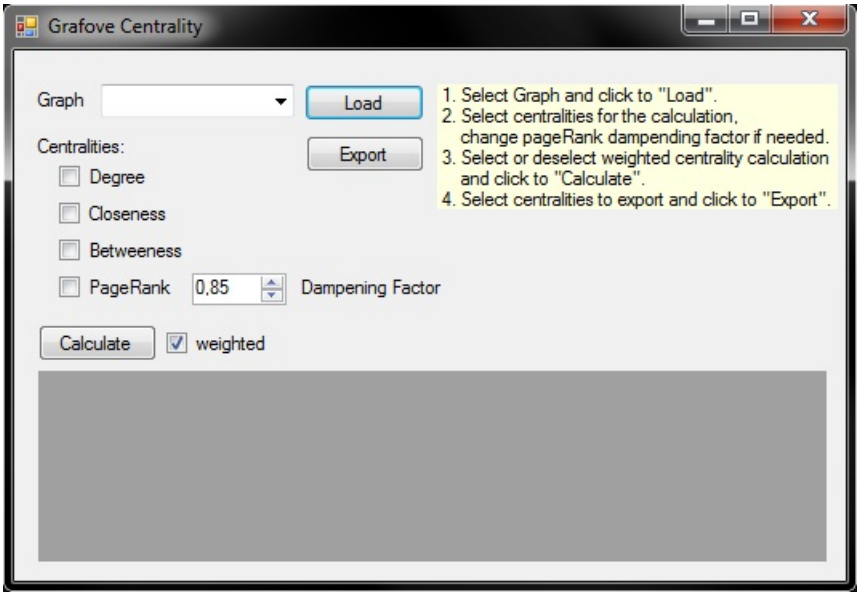
4. Podle potřeby označit nebo odznačit výpočet s ohodnocenými hranami grafu.



Obrázek 7: Diagram tříd



Obrázek 8: Model tříd



Obrázek 9: Uživatelské rozhraní aplikace

5. Stisknout tlačítko "Calculate", díky kterému se provede výpočet označených centralit. Pokud výpočet proběhne v pořádku, vypíše se do info.labelu doba trvání výpočtu a pod ní se opět vypíšou informace o datové kolekci.

6. Posledním krokem je zapsání dat do souboru, to se provede stisknutím tlačítka "Export". Pokud vše proběhne jak má, tak se do souboru zapíšou jen označené centrality a do info.labelu se vypíše oznámení o úspěšném zápisu do souboru.

4 Experimenty

V této kapitole uvádím experimenty prováděné nad kolekcemi dat. Každá kolekce dat je zde popsaná a má uvedený svůj zdroj. Také je ke každé kolekci přiřazen obrázek celé sítě a tabulka zachycující minimální a maximální hodnoty všech centralit pro ohodnocené i neohodnocené hrany. K vytvoření obrázků sítě byl použit nástroj Microsoft Office Excel 2010 doplněný o plugin NodeXL: Network Overview, Discovery and Exploration for Excel. Pro každou kolekci dat a pro každý typ centrality je vytvořen XY-graf znázorňující hodnoty centralit jako křivku. Všechny hodnoty PageRank centrality jsou počítány s hodnotou tlumícího faktoru 0,85. Na konci kapitoly je uvedena tabulka zachycující pro každou ohodnocenou i neohodnocenou datovou kolekci doby trvání výpočtů všech centralit.

4.1 Football

Kolekce *Football* [13], je orientovaná, ohodnocená komplexní síť. Představuje 22 fotbalových týmů, které se účastnily fotbalového Mistroství světa v Paříži v roce 1998. Hráči národního týmu mají často herní smlouvy v zahraničí. To vytváří trh hráčů, kde mohou národní týmy posílat své hráče hrát do zahraničních zemí. Hráči všech 22 národních týmů mají tedy herní smlouvy celkem v 35 zemích.

Vrcholy grafu zde reprezentují jednotlivé země a hrany grafu reprezentují situaci, kdy hráči z jedné země mají uzavřenou herní smlouvu jiné zemi. Váha hrany grafu pak představuje počet hráčů vyslaných z jedné země do druhé.

Graf této kolekce je velmi nesymetrický, protože některé země pouze posílají své hráče hrát do cizích zemí a jiné země naopak pouze přijímají hráče z cizích zemí.

Na obrázku 10 lze vidět grafické zobrazení datové kolekce Football. Tato datová kolekce má celkem 35 vrcholů a 118 hran. Minimální a maximální hodnoty všech centralit vypočítané pro tuto ohodnocenou i neohodnocenou datovou kolekci jsou zaznamenány v tabulce 1.

Na obrázcích 11, 12, 13 a 14 jsou zobrazeny grafy sítě Football, kde na ose X je pokaždé kolekce všech vrcholů grafu a na ose Y jsou hodnoty dané centrality.

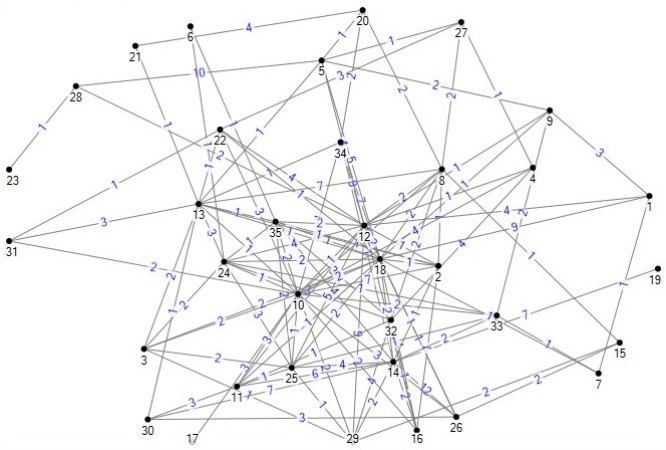
4.2 Les Miserables

Kolekce *Les Miserables* [14], je neorientovaná ohodnocená síť postav, které společně vystupují v románu Victora Huga "Les Miserables". Každý vrchol grafu reprezentuje jednu postavu. Hrany grafu reprezentují dvě postavy, které se společně objevily v jedné kapitole románu. Váha hrany grafu představuje počet těchto společných výskytů dvou postav.

Graf je tvořen celkem 76 vrcholy a 254 hranami.

Na obrázku 15 je zobrazená grafická podoba datové kolekce Les Miserables.

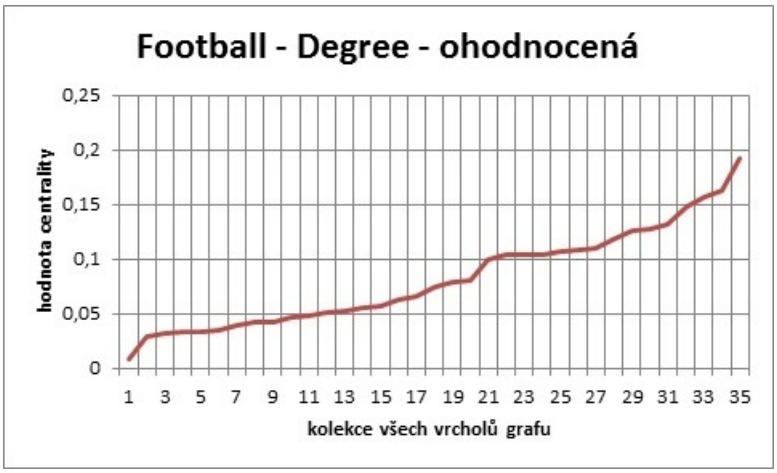
Tabulka 2 zaznamenává minimální a maximální hodnoty všech centralit počítaných pro ohodnocenou i neohodnocenou variantu této datové kolekce.



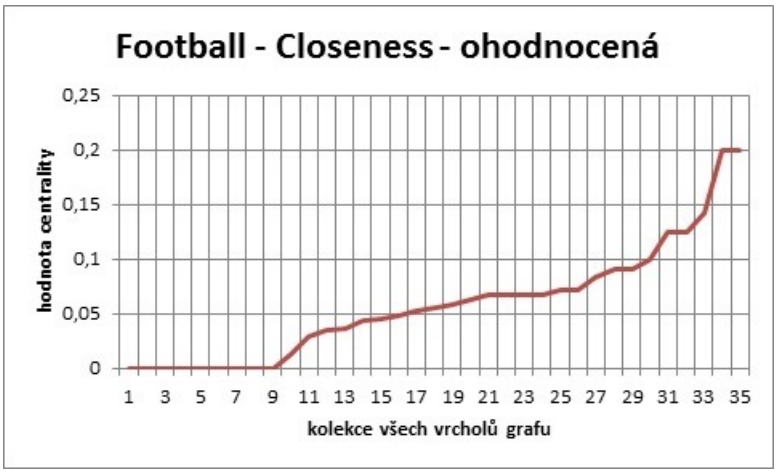
Obrázek 10: Datová kolekce Fotball

Centralita	Ohodnocená	Neohodnocená
min Degree	0,009	0,0294
max Degree	0,1925	0,2941
min Closeness	0	0
max Closeness	0,2	1
min Betweenness	0	0
max Betweenness	20	17
min PageRank	0.65	0.65
max PageRank	3.2336	2.6217

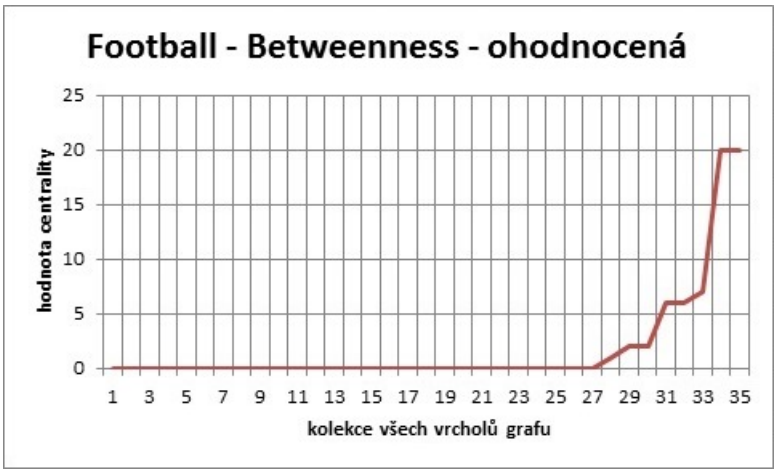
Tabulka 1: Minimální a maximální hodnoty centralit pro datovou kolekci Football



Obrázek 11: Datová kolekce Football - ohodnocená Degree centralita



Obrázek 12: Datová kolekce Football - ohodnocená Closeness centralita

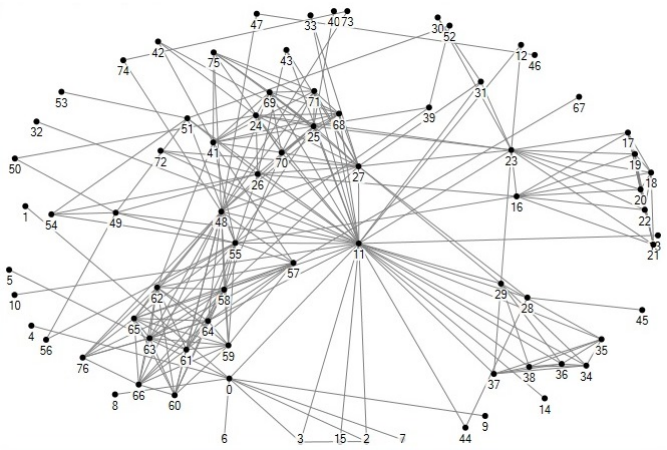


Obrázek 13: Datová kolekce Football - ohodnocená Betweenness centralita

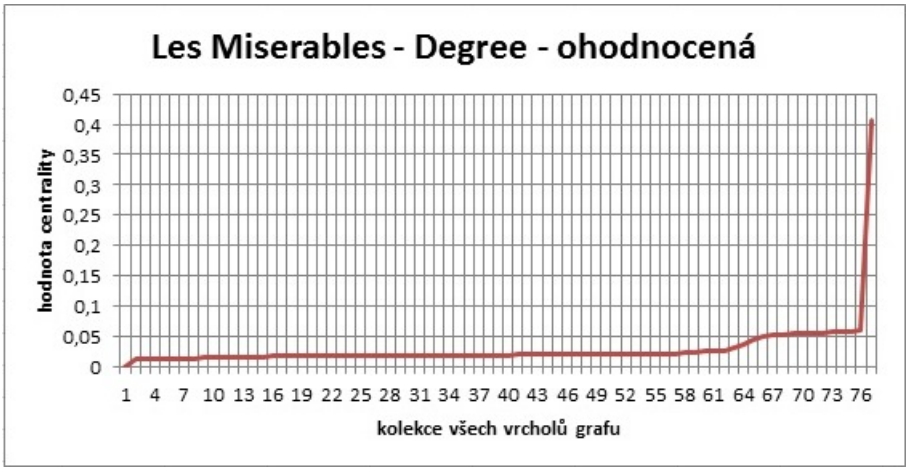
Centralita	Ohodnocená	Neohodnocená
min Degree	0	0
max Degree	0,4079	0.1316
min Closeness	0	0
max Closeness	1	1
min Betweenness	0	0
max Betweenness	412	383
min PageRank	0.65	0.65
max PageRank	8,0137	8.3821

Tabulka 2: Minimální a maximální hodnoty centralit pro datovou kolekci Les Miserables

Obrázek 14: Datová kolekce Football - ohodnocená PageRank centralita



Obrázek 15: Datová kolekce Les Miserables



Obrázek 16: Datová kolekce Les Miserables - ohodnocená Degree centralita

Centralita	Ohodnocená	Neohodnocená
min Degree	0	0
max Degree	0.002	0.002
min Closeness	0	0
max Closeness	0.0556	1
min Betweenness	0	0
max Betweenness	2617	2410
min PageRank	0.65	0.65
max PageRank	60.5523	70.3612

Tabulka 3: Minimální a maximální hodnoty centralit pro datovou kolekci US-Airport

Na obrázcích 16, 17, 18 a 19 jsou zobrazeny grafy sítě Les Miserables, kde na ose X je pokaždé kolekce všech vrcholů grafu a na ose Y jsou hodnoty dané centrality.

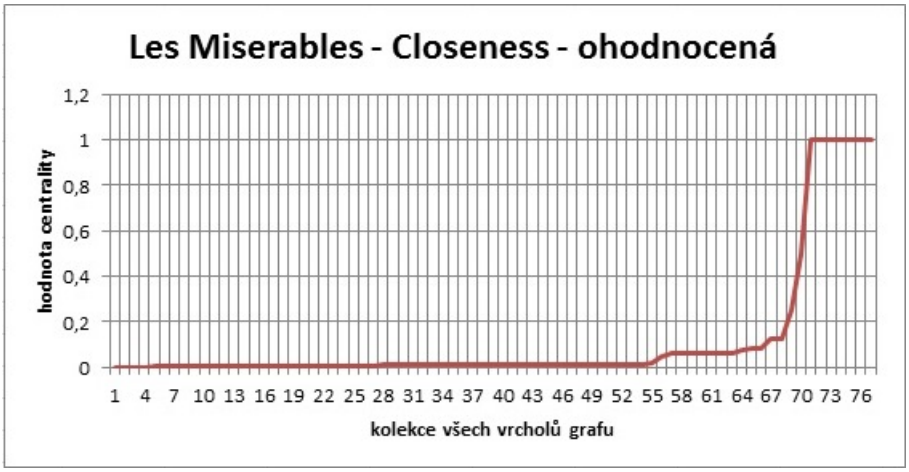
4.3 US-Airport

Kolekce US-Airport, [15], je neorientovaná, ohodnocená síť obsahující data 500 nejvíce vytížených amerických letišť. Každý vrchol grafu představuje jedno letiště a každá hrana grafu představuje letecké spojení mezi dvěmi letišti. Váhy hran pak znamenají počet dostupných míst pro cestující daného leteckého spojení na jeden rok.

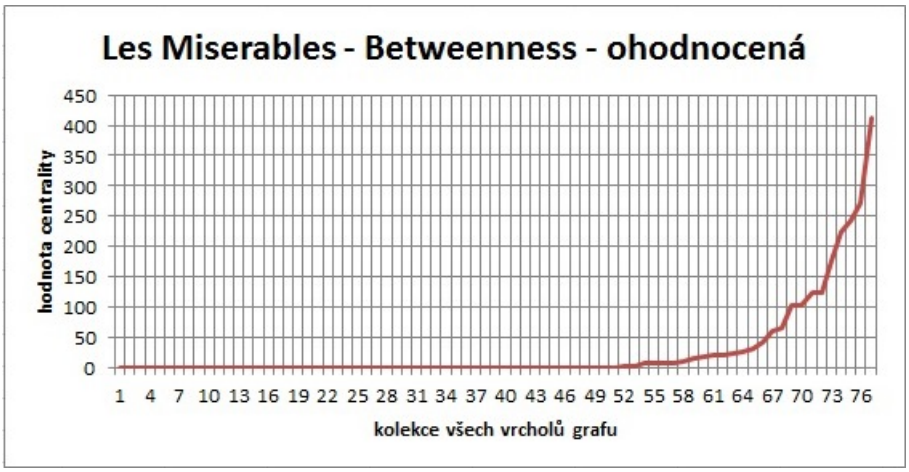
Na obrázku 20 je graficky zobrazena síť letišť datové kolekce US-Airport. Do obrázku sítě nejsou zaznamenány Id jednotlivých vrcholů, ani váhy hran, protože kvůli jejich velkému množství by byl graf velmi nepřehledný.

V tabulce 3 jsou zaznamenány minimální a maximální hodnoty všech centralit pro ohodnocenou i neohodnocenou variantu datové kolekce US-Airport.

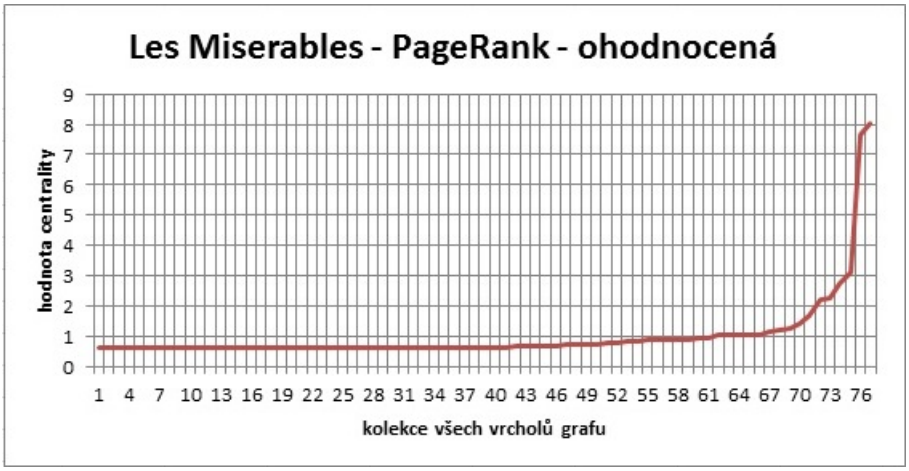
Na obrázcích 21, 22, 23 a 24 jsou zobrazeny grafy sítě US-Airport, kde na ose X je pokaždé kolekce všech vrcholů grafu a na ose Y jsou hodnoty dané centrality.



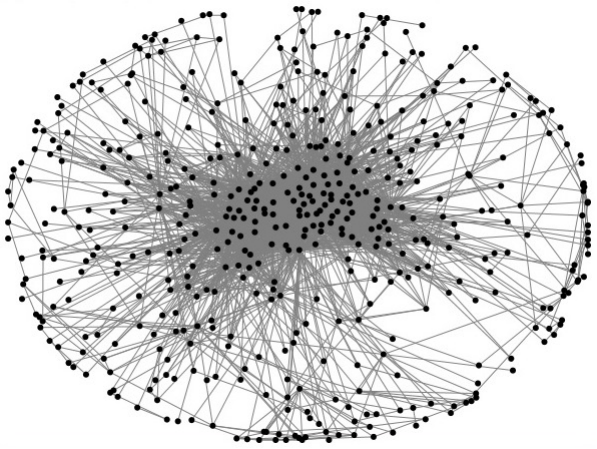
Obrázek 17: Datová kolekce Les Miserables - ohodnocená Closeness centralita



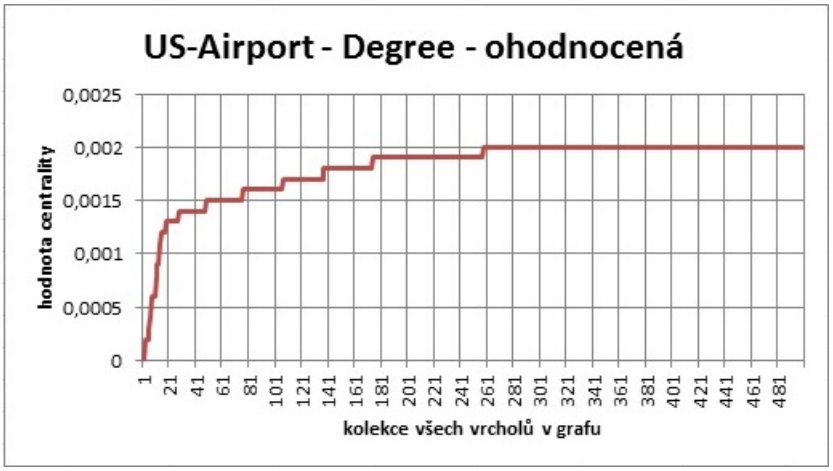
Obrázek 18: Datová kolekce Les Miserables - ohodnocená Betweenness centralita



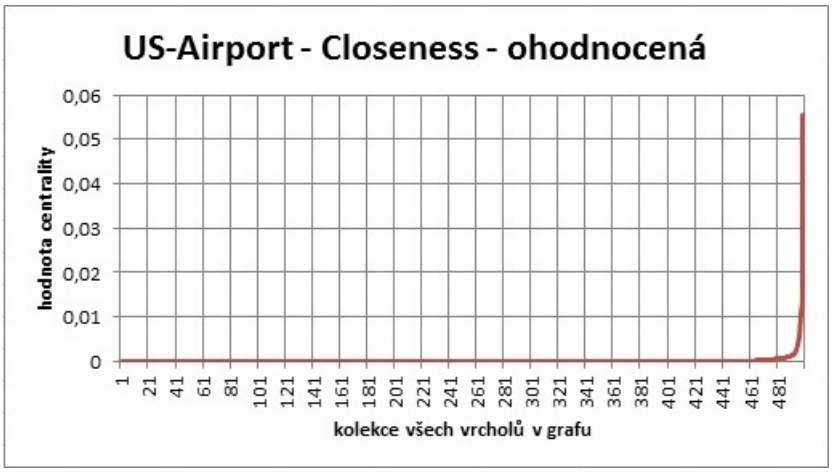
Obrázek 19: Datová kolekce Les Miserables - ohodnocená PageRank centralita



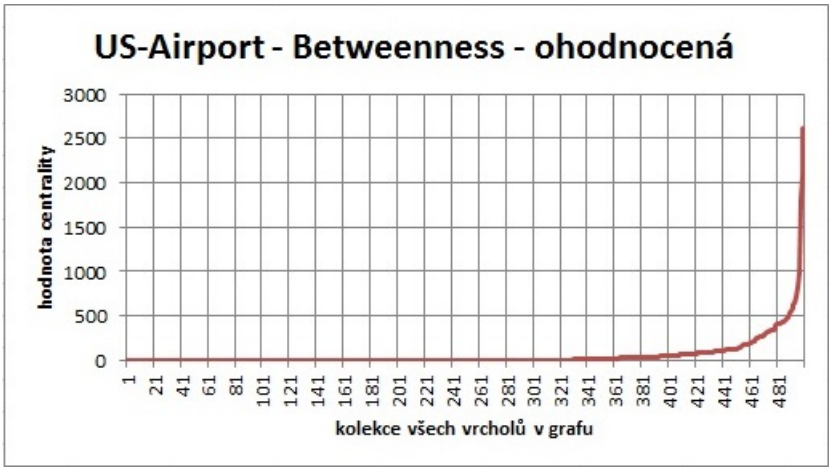
Obrázek 20: Datová kolekce US-Airport



Obrázek 21: Datová kolekce US-Airport - ohodnocená Degree centralita

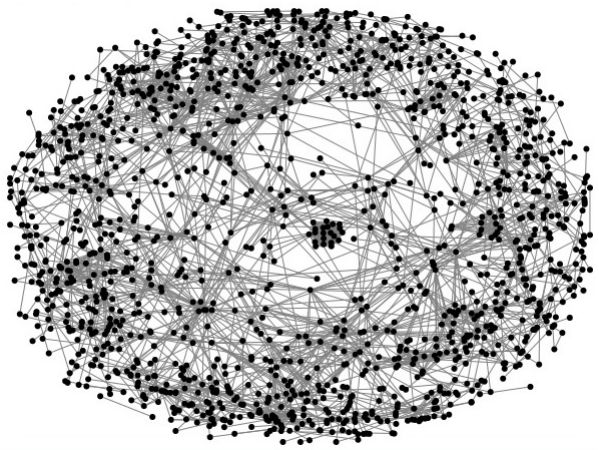


Obrázek 22: Datová kolekce US-Airport - ohodnocená Closeness centralita



Obrázek 23: Datová kolekce US-Airport - ohodnocená Betweenness centralita

Obrázek 24: Datová kolekce US-Airport - ohodnocená PageRank centralita



Obrázek 25: Datová kolekce Network Science

Centralita	Ohodnocená	Neohodnocená
min Degree	0.0003	0.0003
max Degree	0.0007	0.0011
min Closeness	0	0
max Closeness	19	1
min Betweenness	0	0
max Betweenness	302	349
min PageRank	0.65	0.65
max PageRank	4.005	3.9264

Tabulka 4: Minimální a maximální hodnoty centralit pro datovou kolekci Network Science

4.4 Network Science

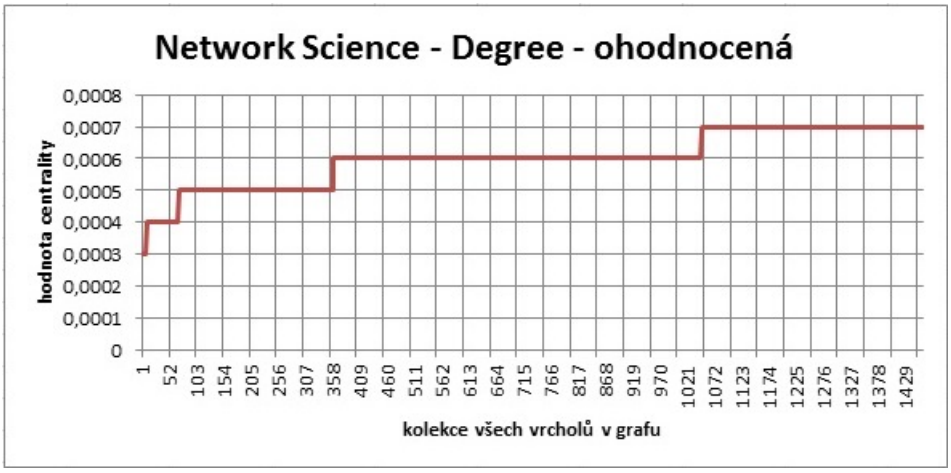
Kolekce Network Science [16], je ohodnocená komplexní síť. Jedná se o spoluautorskou síť vědců pracujících na teorii o sítích. Datová kolekce byla zkompletovaná M. Newmannem v Květnu 2006. Tato komplexní síť byla složena z bibliografie dvou vědeckých článků o sítích, M. E. J. Newman, SIAM Review 45, 167-256 (2003) a S. Boccaletti et al., Physics Reports 424, 175-308 (2006). Datová kolekce se skládá z 1589 vrcholů, tento počet vrcholů odpovídá počtu vědců spolupracujících na teorii o sítích. Váhy hran jsou založeny na počtu společných prací a počtu autorů těchto prací.

Na obrázku 25 je zobrazen graf Network Science.

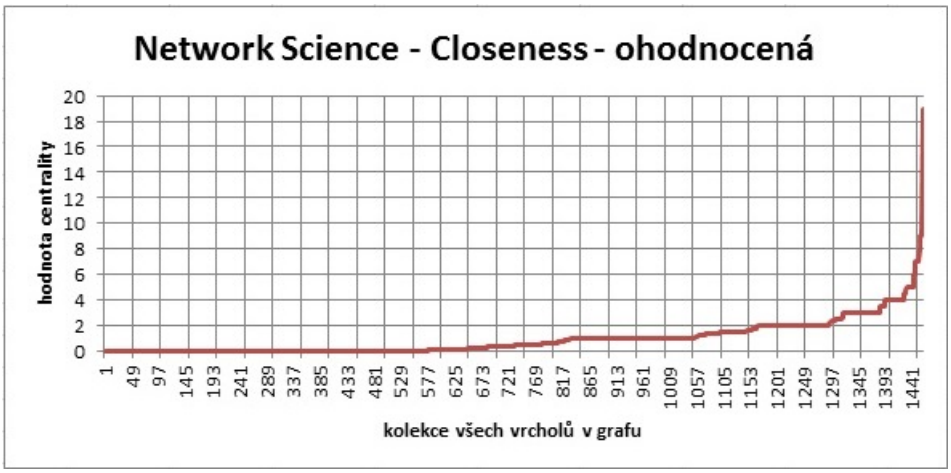
Tabulka 4 zaznamenává minimální a maximální hodnoty všech počítaných centralit pro ohodnocenou i neohodnocenou variantu datové kolekce Network Science.

Na obrázcích 26, 27, 28 a 29 jsou zobrazeny grafy sítě US-Airport, kde na ose X je pokaždé kolekce všech vrcholů grafu a na ose Y jsou hodnoty dané centrality.

Tabulka 5 zaznamenává časovou náročnost výpočtů všech centralit pro ohodnocené i neohodnocené varianty datových kolekcí.



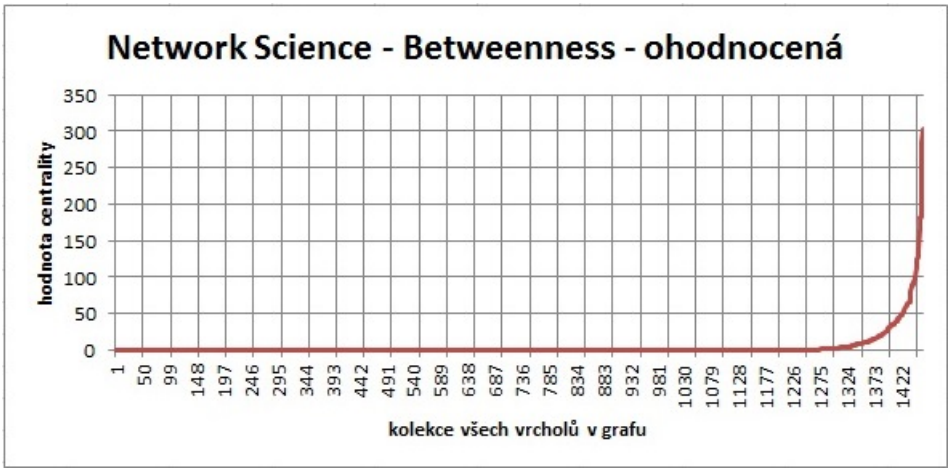
Obrázek 26: Datová kolekce Network Science - ohodnocená Degree centralita



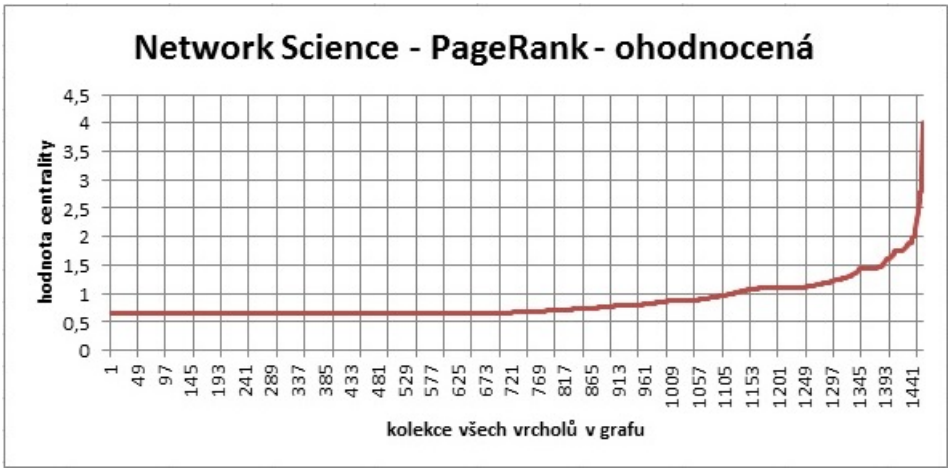
Obrázek 27: Datová kolekce Network Science - ohodnocená Closeness centralita

Datová kolekce	Degree	Closeness	Betweenness	PageRank
Football - ohodnocená	1.8131 ms	8.2656 ms	3.8517 ms	3.1336 ms
Football - neohodnocená	2.3192 ms	8.3911 ms	4.3071 ms	3.2753 ms
Les Miserables - ohodnocená	2.2267 ms	2161.3274 ms	21.0118 ms	5.0355 ms
Les Miserables - neohodnocená	2.7742 ms	2012.4233 ms	21.0237 ms	4.8800 ms
US Airport - ohodnocená	29.4710 ms	2381.3197 ms	752.3575 ms	121.7492 ms
US Airport - neohodnocená	24.9027 ms	2074.4169 ms	740.7351 ms	122.3413 ms
Network Science - ohodnocená	56.0349 ms	23.1763 s	22.7009 s	133.4307 ms
Network Science - neohodnocená	56.5705 ms	23.1885 s	22.8221 s	137.1891 ms

Tabulka 5: Doba výpočtu centralit



Obrázek 28: Datová kolekce Network Science - ohodnocená Betweenness centralita



Obrázek 29: Datová kolekce Network Science - ohodnocená PageRank centralita

5 Závěr

Cílem této práce bylo navrhnout a naimplementovat aplikaci schopnou načíst kolekce dat ze souboru, spočítat čtyři typy centralit a vypočtená data zapsat do souboru.

Podařilo se načíst různé typy ohodnocených a orientovaných komplexnícxh sítí, spočítat ohodnocené i neohodnocené varianty centralit jejich vrcholů a výsledky zapsat do souboru.

Pro nalezení nejkratších cest mezi všemi dvojicemi vrcholů v grafu byl původně v aplikaci naimplementovaný Dijkstrův algoritmus. Tato implementace způsobila, že během provádění experimentů byl výpočet Closeness a Betweenness centrality časově velmi náročný. Někdy trval výpočet u velké kolekce dat (1000 vrcholů a více) až minutu. Proto byl v aplikaci Dijkstrův algoritmus nahrazen Floyd-Warshallovým algoritmem, který je implementačně jednodušší.

Aplikace by mohla být ještě doplněna o implementace algoritmů pro výpočet dalších typů centralit.

6 Reference

[1] NEWMANN M., Networks: An Introduction, 1. vyd., Oxford University Press, Inc., New York, NY, USA, 2010, ISBN-10: 0199206651, ISBN-13: 978-0199206650.

[2] PELÁNEK R.. Modelování komplexnívh sítí. Nakladatelství Masarykovy univerzity, 2011. ISBN 978-80-210-5318-2.

[3] FREEMAN L. C., A Set of Measures of Centrality Based on Betweenness, 1977, [online], [cit. 2015-04-29]. Dostupné z <http://moreno.ss.uci.edu/23.pdf>.

[4] FREEMAN L. C., Centrality in Social Networks Conceptual Clarification, 1979, [online], [cit. 2015-04-29]. Dostupné z <http://moreno.ss.uci.edu/27.pdf>.

[5] FERRARA E., Mining and Analysis of Online Social Networks, 2012, [online], [cit. 2015-04-29]. Dostupné z <http://www.emilio.ferrara.name/wp-content/uploads/2011/06/thesis.pdf>.

[6] HANNEMAN R. A. a RIDDLE M., Introduction to Social Network Methods, 2005, [online], [cit. 2015-04-30]. Dostupné z <http://faculty.ucr.edu/~hanneman/nettext/>.

[7] LANGVILLE A. N. a MEYER C. D., Introduction to Web Search Engines, 2006, [online], [cit. 2015-04-30]. Dostupné z <http://press.princeton.edu/chapters/s8216.pdf>.

[8] International Journal of Soft Computing and Engineering (IJSCE). Weighted Page rank Algorithm Based on Number of Visits of Links of Web Page [online], poslední revize Červenec 2012 [cit. 2015-04-10]. Dostupné z <http://www.ijsce.org/attachments/File/v2i3/C0796062312.pdf>.

[9] OCHODKOVÁ E., Grafové algoritmy a komplexní sítě [online], [cit. 2015-04-05]. Dostupné z <http://www.cs.vsb.cz/ochodkova/courses/gaks/gaks1_2014.pdf>.

[10] KOVÁŘ P., Teorie grafů, [online], [cit. 2015-04-31]. Dostupné z <http://homel.vsb.cz/ kov16/files/skriptum_teorie_grafu_rozsirene_tisk.pdf>.

[11] NYKL M., Určování významnosti vrcholů grafu: PageRank a jeho modifikace [online], [cit. 2015-04-29]. Dostupné z <http://www.kiv.zcu.cz/site/documents/verejne/vyzkum/publikace/technicke-zpravy/2013/tr-2013-09.pdf>.

[12] OPSAHL T., Weighted Networks, [online], [cit. 2015-04-15]. Dostupné z <http://toreopsahl.com/tnet/weighted-networks/>.

[13] Datová kolekce, football [online], [cit. 2015-03-04]. Dostupné z <http://vlado.fmf.uni-lj.si/pub/networks/data/sport/football.htm>.

[14] Datová kolekce, Les Miserables [online], [cit. 2015-03-04]. Dostupné z <http://www-personal.umich.edu/~mejn/netdata/>.

[15] Datová kolekce, US Airport [online], [cit. 2015-03-04]. Dostupné z <https://sites.google.com/site/cxnets/usairtransportationnetwork>.

[16] Datová kolekce, Network Science [online], [cit. 2015-03-04]. Dostupné z <http://nexus.igraph.org/api/dataset_info?id=21&format=html>.

[17] Obrázek sociální sítě, [online], [cit. 2015-04-05]. Dostupné z <http://tech18.com/global-map-social-networking-infographic.html>.

[18] Obrázek internetové sítě, [online], [cit. 2015-04-05]. Dostupné z <http://spacecollective.org/meganmay/1024/PROPOSAL-FOR-A-NEW-SOCIETY-call-to-action>.

[19] Obrázek obchodní sítě, [online], [cit. 2015-04-05]. Dostupné z <http://www.akitarescueoftulsa.com/networking-diagram/>.

[20] Obrázek sítě letových tras, [online], [cit. 2015-04-05]. Dostupné z <http://www.zib.de/features/mathematics-traffic-and-transport>.

[21] Obrázek sítě neuronů mozku, [online], [cit. 2015-04-05]. Dostupné z <http://www.pnas.org/content/109/15/5549/F4.expansion.html>.